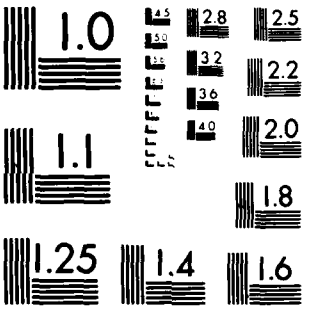
AD-A134 585 MEASURING THE READABILITY OF TRAINING MATERIALS BY THE 1/1
PLATO IV COMPUTER-..(U) SCHOOL OF APPLIED AEROSPACE
SCIENCES CHANUTE AFB IL G P SCHARF 15 SEP 74 CH-74-604
UNCLASSIFIED SBI-AD-F630 042 F/G 5/9 NL

END
DATE FILMED
11 83
DTIC



CHANUTE PROJECT REPORT 74-604

MEASURING THE READABILITY OF TRAINING MATERIALS BY THE PLATO IV COMPUTER-BASED INSTRUCTIONAL SYSTEM

15 September 1974

DEPARTMENT OF THE AIR FORCE
3330TH TECHNICAL TRAINING WING (ATC)
CHANUTE AIR FORCE BASE, IL 61868

REPLY TO ATTN OF TTGH

18 AUG 1983

SUBJECT Request for Scientific and Technical Reports

TO Administrator
Defense Technical Information Center
ATTN: DTIC-DDAB P83-0290
Cameron Sta BG 5
Alexandria VA 22314

1. I regret that it took so long to respond to your request. I was unable to locate the document at Chanute. A copy was finally located at our Headquarters' Tech Advisory Service Branch.

2. Notice that page 1 is missing from the document. As soon as I receive a copy I will send it to your office.

Paul C. Aschenbrenner

PAUL C. ASCHENBRENNER
Chief, PLATO Training Branch

2 Atch
1. DTIC-DDAB P83-0290 Ltr, 18 Jul 83
2. CH 74-604

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

ACKNOWLEDGMENT

I wish to acknowledge the assistance of and extend my thanks to the PLATO IV authors, the typists, the technical editors, and the training specialists who participated in the experiment discussed in this report. Special appreciation goes to MSgt [illegible] C. Dennis, SSgt (now Mr.) Dennis W. Mitts, and SSgt (now Mr.) [illegible] W. Bourne for their development of [illegible] PLATO IV [illegible] reading [illegible]. Without their interest and [illegible] participation, the project would not have been possible.

DEPARTMENT OF THE AIR FORCE
USAF School of Applied Aerospace Sciences
Chanute Air Force Base, Illinois 61868

CHANUTE PROJECT
REPORT 74-602
15 September 1974

MEASURING THE READABILITY OF TRAINING MATERIALS BY THE PLATO IV COMPUTER-BASED INSTRUCTIONAL SYSTEM

ABSTRACT

This report discusses an experiment in which the PLATO IV computer-based instructional system was programmed to measure the reading grade levels of written training materials. The lesson in the PLATO system that performs this task is called PIRL (pronounced pearl), which stands for PLATO Indicated Reading Level. In this Air Force experiment, samples of technical training materials from five different career fields were checked for their grade level by 24 technical writers using the Fog count system. The same samples were then checked by nine education/training specialists using the Flesch system. The samples were then typed into the PLATO IV system by six clerk/typists and then by seven PLATO IV authors.

While the results showed some variation in grade levels within all four groups, the mean grade levels of each sample for all groups were quite close. An analysis of variance of the all-sample means of the four groups indicates no statistically significant differences among them. It is concluded that the PLATO IV lesson PIRL is an easily used and valid means for determining and expressing the difficulty level of written materials as a grade level.

George P. Scharf, GS-13
Chief, Msl & Sp Plans Section

This report has been reviewed and is approved.

Clinton D. Gillespie

CLINTON D. GILLESPIE, Colonel, USAF
Commander

TABLE OF CONTENTS

instruction system to Chanute. The terminals are supporting a service test to see just how computer-assisted instruction might be used in teaching a technical training course. The project is monitored at Chanute by the Training Research Applications Branch. Because branch personnel had been interested in the problem of the reading difficulty of training literature for several years, it was proposed that PLATO IV be investigated for possibilities for automating the determination of reading grade levels of training materials. The remainder of this paper will discuss the results of this investigation.

SECTION C - METHOD

6. In deciding the ground rules for programming lesson PIRL in the PLATO system to determine reading difficulty levels, the experiences with past systems were used as the basis for the program parameters.

a. The PLATO IV terminal uses a keyboard similar to that of a regular typewriter. The first ground rule set up was that the readability determination system should require no variations from normal typing procedure. This ground rule has been modified to the extent that no periods are used in a sample being checked except to indicate the end of a sentence. Abbreviations and terms such as etc., i.e., and e.g. are typed into the system without the period.

b. Another rule adopted was that numbers would not be typed into the system; e.g., $1,000,000.

c. The third ground rule established was that the PIRL readout should be in grade level rather than on some other scale. In previous experiments, we have found that our technical writers (as well as training specialists and supervisors) can interpret a grade level, but do not relate well to a numerical scale.

7. In developing a formula for the PLATO system to use to determine grade level, the first consideration was to establish that the factors used relate to those found in other reading level indices. The element relating to sentence length (words per sentence) is identical to that found in most currently used indices such as the Flesch and Fry Formulas; so, no further verification for using this component in the PIRL formula was considered necessary. The verification of the relationship within the word structure was also virtually self-evident since an association does exist between the number of letters in a word and the number of syllables it contains. Also, the average number of letters per word apparently bears a fairly close relationship to the proportion of words included in a list of most common words, which is the basis of the Dale-

Chall readability formula. Consequently, the two factors used in the PIRL formula (average number of letters per word and average number of words per sentence) should provide a readability index valid with other systems if proper weightings for each factor are determined.

a. To obtain a grade level reading, samples to be measured are selected and the material is typed into the PLATO lesson called PIRL. Detailed instructions on how to use this lesson are included in the program. However, these are not called up unless the user indicates he needs them by pushing the "Help" key. The first screen display seen when signing into PIRL is shown in Figure 1. After a sample has been typed into the PIRL lesson, the "Next" key is pushed. The computer then begins inspecting the words and assigns a point value to each word as follows:

1, 2 or 3 letters	1 point
4, 5 or 6 letters	2 points
7, 8 or 9 letters	3 points
10 or more letters	4 points

b. Following the word count, the computer counts the number of complete thoughts/sentences as indicated by a period, question mark, exclamation mark, colon, or semi-colon. The total number of word points is then divided by the number of sentences. This answer is then divided by three, which is then displayed on the screen as the PIRL, which is indicated as a grade level. An illustration of how this appears on the screen may be seen in Figure 2.

c. If the "Shift" "Next" keys are pressed, the material on the screen is erased and a new sample may be typed into the system. However, if a sample is typed into the system and the typist wishes to save the information concerning the sample, the "Data" key is pressed. A data collection file is then started and information on this first sample then appears on the screen. When the "Next" key is pressed, this data is saved while another sample is typed into the system. Data for up to 15 samples may be held in this way, with the mean of all the samples automatically calculated whenever a new sample is added. A data collection file screen display may be seen in Figure 3. This file not only shows the grade level of each sample and the mean of the samples, but also shows the sentence count, total word count, and the numbers of words with 1, 2, or 3 letters; 4, 5, or 6 letters; 7, 8, or 9 letters; and 10 or more letters.

d. The appropriateness of the weightings given to word lengths was established by correlating grade levels calculated by PIRL with grade

-HELP- if you need instructions to use this lesson

P.I.R.L.

Plato Indicated Reading Level

TWO GROUND RULES:

1. Omit ALL periods EXCEPT at the end of sentences.

2. Numbers DONT'T count, so omit them.
   (numbers = numerics i.e..3,107,1t etc.)

-NEXT-

Figure 1. First Display Seen When Signing Into Lesson PIRL

TYPE MATERIAL AT THE ARROW: Press -NEXT- when finished.
-HELP-

> This report discusses an experiment in which the PLATO computer based instructional system was programmed to measure the reading (grade) levels of written training materials. The lesson in the PLATO system that performs this task is called PIRL, which stands for PLATO

Indicated Reading Level. In this Air Force experiment, samples of technical training materials from five different career fields were checked for their grade level by 11 technical writers using the fog count system. The same samples were then checked by nine education/training specialists

using the Flesch system. The samples were then typed in the PLATO IV system by six clerk typists and then by seven PLATO authors. An analysis of variance of the all-sample means in the four groups indicates no statistical significant difference among them. It is concluded that

the PLATO IV lesson PIRL is an easily used and valid means for determining and expressing the difficulty level of written materials as a grade level.

Reading level...15.3
Number of sentences.....7 Total number of words...158
-SHIFT NEXT- for new sample, -DATA- to store this data
Time... 7.7 minutes

Figure 2. Screen Display of Sample and Its Data

to have room for 10 more samples before
the data collection file is filled.

(458)	(%d)	(1.*)	#words	*sem	FIRL
.2	.2	8	7	3	1. 3
-	8	2	1	13	1 4 7
.	5	-	5	11	2 5
.	1	7	6	15	1 1
11	19	8	4	43	1 1
16	41	25	23	116	7 ...TOTAL
5	8	5	4	23	1 ...MEAN

Year FIRL of 5 samples... 1.*

.E T- TO STORE ... DATA
-SHIFT DATA- TO CALCULATE THIS
C TO END ENTRY ... NEW SAMPLE

Figure 3. Screen Display of Data Collection File of Five Samples

levels of the same samples established by other systems. Thus, the PIRL formula, like most other readability formulas, is derived from ratios representing word difficulty (number of letters per word) and sentence difficulty (number of words per sentence).

5. To obtain a relatively broad base for verifying the PIRL formula, initial work was done with five, 2,000 word samples. Grade levels were determined by the Flesch count, the fog count, and by PIRL and found to have a high degree of correlation. Then, to verify the PIRL lesson in an operational situation, the following experimental design was developed.

a. Five 250 word samples of random paragraphs from existing technical training materials were selected. Each sample was from an entirely different career field; i.e., weather, missiles, pneudraulics, jet engines, and automotive.

b. Twenty-seven technician/authors from the five training departments were asked to perform a fog count of the samples and record their findings. However, no technician performed a fog count on materials originated in his department. As it turned out, there were 21 author/technicians who did a fog count on each sample.

c. Six typists, unfamiliar with the PLATO system, were asked to type the five samples into the PLATO system and to determine and record the grade level of each sample.

d. Seven PLATO authors were asked to type the five samples into the PLATO system and to determine and record the grade level of each sample.

e. Nine Education and Training Specialists with curricula experience were asked to perform a Flesch count of the five samples. This count was then converted to its grade level using the table in Flesch's book, How to Test Readability.

SECTION D - RESULTS AND DISCUSSION

6. The results achieved by the four groups of people used in this study are shown in Tables 1 thru 4.

a. Fog count results are shown in Table 1. Differences in the grade levels obtained by different writers can be easily seen in the range for each sample.

Sample	Grade Level Range	Mean Grade Level	Standard Deviation
I	11-14 = 3	13.2	1.0
II	8-11.2 = 3.2	9.4	1.2
III	12.1-19.5 = 7.4	15.5	1.7
IV	8.3-11.7 = 3.4	9.8	1.0
V	13.5-16.0 = 2.5	14.6	0.9

Table 1. Results of the Manual Fog Count of Five Samples by 21 Technician/Authors.

b. The indicated reading levels as determined by the typists using lesson PIRL are shown in Table 2. Because typing is basically a mechanical procedure, one would expect a very high correlation of the typists'

Sample	Grade Level Range	Mean Grade Level	Standard Deviation
I	12.7-13.8 = 1.1	13.5	.3
II	9.3-10 = .7	9.5	.2
III	14.4-14.9 = .5	14.8	.1
IV	10.7-11.2 = .5	10.9	.17
V	14.2-14.6 = .4	14.5	.1

Table 2. Results of PIRL Grade Level Determinations of Five Samples by Six Clerk/Typists.

determination of reading levels. However, the PIRL method does depend upon typing accuracy, and with the relatively small samples used in this experiment, typing errors did cause noticeable grade level variations. The typing errors, in most cases, were probably caused by the different touch of the PLATO terminal keyboard.

c. Results achieved by the PLATO authors using lesson PIRL are shown in Table 3.

Sample	Grade Level Range	Mean Grade Level	Standard Deviation
I	[illegible]-15.2 = 1.[illegible]	[illegible]	.5[illegible]
II	[illegible].3-9.[illegible] = .[illegible]	[illegible]	.0[illegible]
III	[illegible].8-17.[illegible] = .[illegible]	[illegible].9	[illegible]
[illegible]	1[illegible].8-10.[illegible] = .[illegible]	[illegible]	.[illegible]
[illegible]	[illegible]-1[illegible].5 = .[illegible]	[illegible]	[illegible]

Table 3. Results of [illegible] Grade Level Determination of Five Samples by Seven PLATO IV Authors.

For the PLATO authors, variations in grade level [illegible] typing errors probably occurred because they are not professional typists. In comparing Tables [illegible] and 3, however, it doesn't take any [illegible] statistical technique to recognize that there [illegible] much [illegible] less grade level variation among the typists using [illegible] than among [illegible] counters. Both the ranges and the standard deviations are much less for the [illegible] system users.

[illegible] the results from nine [illegible] and training specialists [illegible] determining grade levels [illegible] the same five samples by the Flesch count. These higher level people were chosen for this part of

Sample	Grade Level Range	Mean Grade Level	Standard Deviation
I	1[illegible].0-1[illegible].0 = 1	[illegible]	[illegible]
II	[illegible].0-9.[illegible] = .[illegible]	[illegible]	.[illegible]
III	1[illegible].0-1[illegible].0 = 1	[illegible]	.[illegible]
IV	1[illegible].0-11.0 = 0	11.[illegible]	.0[illegible]
V	1[illegible].0-1[illegible].0 = [illegible]	1[illegible]	.[illegible]

Table 4. Results of [illegible] of Five Samples by Nine Education and Training Specialists.

the study because of their background and experience in curricula, their educational level, and the fact they should be more skilled at identifying syllables, which is the heart of the Flesch system. Even higher level people are not infallible, and in one 500-word sample, there was a count variation of 40 syllables between the highest and lowest counts. It is also important that one be aware that the Flesch system has a range rather than one value to convert to grade levels. For example, a Flesch raw score of 42-46 approximates a grade level of 12. The range from 47-51 is a grade level of approximately 11. The range of 52-57 is approximately 11.5. It is this range factor that accounts for the fact that for sample IV every person using the Flesch system came up with a grade level of 11.0 for the sample, even though there were variations in their syllable counts.

10. Possibly the most significant table for validating the PIRL formula, however, is Table 5. Here is shown the mean of all the five samples added together for each method. In other words, this is the mean grade

Mean Grade Level for 21 Technician/Writers Using the "fog count" method	12.50
Mean Grade Level for Six Clerk/Typists Using the PLATO IV lesson called PIRL	11.61
Mean Grade Level for Seven PLATO IV Authors Using the lesson called PIRL	12.0
Mean Grade Level for Nine Education and Training Specialists using the "Flesch" count	12.35

Table 5. Mean of Grade Levels for the Five Test Samples Treated as One, Large Sample.

level of an equivalent 2,000 word sample. As can be seen, the means for all methods used to determine grade levels are quite close. An analysis of the variance of the grade levels found by the four groups is shown in Table 6, and indicates that there is no statistically significant difference in the grade level determinations. All computations for this study were performed using statistical packages available in the PLATO IV system.

F = .312	d.f. 3,39	P = 0.8169

Table 6. Analysis of Variance of the Grade Levels of 2,000 Word Samples by Four Groups.

11. Exact figures were not always kept on the amount of time it took the different people to do their grade level determinations. About half-way through the experiment a timing system was programmed into PIRL. Now, when the PIRL readout is given, a report on the amount of time it took to type in the sample will also appear on the terminal screen (see Figure 2). Both estimates and actual times the participants in this experiment took to determine the reading levels of the five samples were used to arrive at the following rough mean figures:

a. Fog counters	b. hours
c. Client Typist using PIRL	d. minutes
e. PLATO authors using PIRL	f. minutes
g. Flesch counters	h. hours

SECTION E - CONCLUSIONS AND RECOMMENDATIONS

12. It should be kept in mind that there are several factors concerning readability which formulas cannot measure. Such things as imagery, subject matter, word order, and idea organization are also quite (if not more) important than word and sentence length in determining readability. In addition, formula scores can be inaccurate due to errors in sampling or in their application. However, difficulty levels based upon word and sentence length do serve as fair indicators of the understandability of the material. High grade levels can alert a writer that there may be a need to rewrite or restate his material in order to improve its chances of being easily understood. This is especially relevant in Air Force technical writing where experience has shown reading aptitudes may be rather low even when the student is a high school graduate and has high mechanical or electronic AQE scores. Thus, readability ratings can and do serve a useful purpose as a tool to help make one's writing clear and more understandable.

13. As a result of the analysis and experience of this study, the lesson PIRL in the PLATO IV system appears to be an easily used and valid means for determining and expressing the readability of written materials at a grade level. There are now PLATO IV terminals at three AFT centers: Chanute, Lowry and Sheppard. It is recommended that training departments and technical writers at these bases take advantage of this and use lesson PIRL to check the reading levels of their training materials.

SECTION F - BIBLIOGRAPHY

1. *AFM 10-4, Guide for Air Force Writing, U.S. Government Printing Office, Washington D.C., 1960.

2. TP-75100-1070, See the Light Before, You Write, CAFB, Illinois, 1950.

3. Dale, E., and Chall, J., "A Formula for Predicting Readability," Educational Research Bulletin, 1948.

4. Flesch, Rudolph, The Art of Readable Writing, Harper and Brothers, N.Y., 1949.

5. Flesch, Rudolph, How to Test Readability, Harper and Brothers, N.Y., 1951.

6. Fry, Edward, "A Readability Formula that Saves Time," Journal of Reading, 11:577, April 1968.

7. Kincaid, J. P., Yasutake, J. Y., and Geiselhart, R., Use of the Automated Readability Index to Improve Comprehensibility of Technical Orders, SEG-TR-67-47, November 1967.

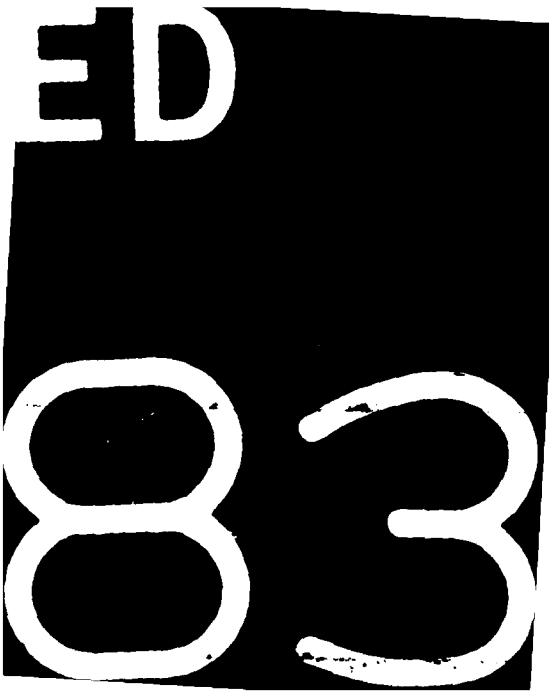
8. Klare, George R., The Measurement of Readability, Iowa State University Press, Ames, Iowa, 1963.

9. Scharf, G. P., A Feasibility Test of the Automated Readability Index Apparatus, ATC PR 69-22, 4 September 1969.

10. Smith, E. A. and Senter, R. J., Automated Readability Index, AMRL-TR-66-220, November 1967.

*Now AFP 13-2, Guide for Air Force Writing. The Fog count method of grade level determination has been deleted from this pamphlet.

The PIRL for this report is 14.8, a reading level of "difficult" in the Flesch system. Less than 30% of the adult population in the United States can read at this level.


ED
83